Sorting the Wheat from the Chaff: A Study of the Detection of Alarms

Neville A. Stanton, David J. Harrison, Karen L. Taylor-Burge & Les J. Porter
Department of Design
Brunel University
Coopers Hill Lane
Egham, Surrey,
TW20 OJZ, UK.

This research in this paper considers the evidence on the success of alarm reduction strategies reported in the open literature. Despite strong beliefs to the contrary, the empirical evidence suggests that alarm reduction strategies have not been as successful as initially expected. This seems to be due to the fact that alarm reduction strategies actually deprive process control operators of information. In order to determine the ability of people to sift through alarm information, a study of alarm detection with three ratios of target to non-target alarms was devised (i.e. 2%, 6%, and 10%) and the information was presented at three rates (i.e. 1 second, 4 seconds and 8 seconds). The results show that the ratio of target alarms has no effect on detection performance, but the temporal rate does. Given that process operators are rarely required to acknowledge alarm information in real-time, it is suggested that more emphasis should be placed on initial definition of alarms and better presentation methods, rather than attempts to block the flow of alarms that have already been triggered.

KEYWORDS: Alarms, False Alarms, Alarm Reduction, Signal Detection.

PROBLEMS WITH ALARMS

The inappropriate presence of alarms can cause substantial problems for the process operator (Woods, O'Brien & Hanes, 1987). Typical problems are; the avalanche of alarms during a major transient or shift in operating mode, standing alarms, alarm inflation, nuisance alarms, and alarms serving as status messages (Stanton & Baber, 1995; Woods et al, 1987; Hoenig, Umbers & Andow, 1982; Andow & Lees, 1974). In a major incident, alarm presentation rate may be somewhere between 50-300 alarms a minute in a nuclear power station (Hickling, 1992). This may lead to problems for the operator in being able to identify and respond to alarms that are worthy of attention. Certainly, the limited number of actions that arise from alarms might suggest that there is a lot of redundant information present, e.g. Kragt & Bonten (1983) found that only 7% resulted in operator action. The problems appear to stem from the design of alarms based on 'normal' operation (Singleton, 1989) and on a 'one measurement - one indication' philosophy of presenting essentially raw plant data (Goodstein, 1985). However, a change in plant state would mean a change in what could be considered 'normal', i.e. what is 'normal' in start-up, maintenance, and shut-down? In addition, the oscillatory behaviour of a variable that is close to its alarm parameter can lead to distrust of the alarm. Hale & Glendon (1987) propose that a shift in confidence occurs, such that the next time the alarm occurs, the first hypothesis the individual will have is that the alarm is a false one. This lack of trust will grow with the number of false alarms experienced.

Sorkin (1989) suggest that individuals are regularly disabling warning systems in locomotive, aircraft and process industries. All the examples are from situations where critical events could arise (e.g. potentially life threatening incidents). Sorkin suggests that alarm systems may be working against individuals, rather than for them, i.e. high alarm rates, aversive signals and false alarms. The attention-getting properties of the alarm should not overwhelm the sensory channels (Hale & Glendon, 1987) and consideration should be given to the 'human-plus-alarm' system (Sorkin, 1989). Andow (1983) suggests that diagnosis is often difficult, and the alarm system does little to help. Computer-based alarm systems have been justified on a number of counts; more flexible control and optimisation of process conditions, providing data of better quality and providing better process and management information (Zwaga & Veldkamp, 1984). Computer-based systems, whilst initially seen as a panacea to the problem, have apparently increased operators' difficulties. This is due to: increased system complexity, provision of even more information, and an increased emphasis on the monitoring task (Hoenig, Umbers & Andow, 1982).

Zwaga & Veldkamp (1984) note that dangerous process conditions can develop with oscillating alarms, as operators tend to acknowledge them prior to determining their location. Once acknowledged, certain types of alarm media make it difficult to determine 'last-up', such as annunciators which go into a steady "alarm-on" state. Whilst this is not a problem if only a small number are present, a large number of alarms make this search task very difficult. Combs & Aghazadeh (1988) argue that serial displays are problematic because they mask alarms, as they build up into a queue and remain unanswered. Certainly there will be a trade off in design between the number of items on a page and the number of levels of pages in the hierarchy. However, Combs & Aghazadeh propose that parallel rather than serial displays provide the solution. They argue that a parallel display could reduce response time, decrease training and increase process continuity. However, they have not substantiated their claims, and it appears that this is largely a return to the kind of philosophy that underlies annunciator panels or plant mimics, which are not without problems.

With annunciator panels or large plant mimics, it can be difficult to detect a new alarm initially if attention is focused on another part of the panel. Once detected and accepted, its status looks the same as any other alarm on the panel so the operator is deprived of sequence of events information. However, with scrolling text displays on VDUs the operator has no trouble observing recent alarms and the order of presentation, but s/he is not provided with any spatial information about the relative location of these events, and earlier alarms may scroll off the screen. Whether operators would use all of the information, even if it were available, is questionable. Andow and Lees (1974) cited Duncan (1972) who showed in a study that when 7 alarms came up simultaneously the skilled operator appeared never to use more than four, and often only used one. Apparently the operator's skill was characterised by using a set of heuristics which enabled a choice from a small set of alternatives, taking high probability paths and checking selected readings. Typically an operator might, upon detecting an abnormal condition, identify the present state and extrapolate future states. Thus, operator diagnostic behaviour has three major elements; historical (identification of problem space), futuristic (extrapolation and prediction of future states) and planning (proposing preventative or corrective action). It is not really possible to isolate alarm systems from the rest of the information display system, and operators rely on both to support their activities.

REDUCTION OF ALARMS

From the arguments presented in the foregoing section it is reasonable to suppose that alarm reduction techniques may alleviate many of the difficulties encountered. Kortlandt & Kragt (1983) surmised that the limited number of actions following an alarm confirms that the main function of the alarm system is as a monitoring tool, i.e. the majority of alarms are not alarms in the sense that a dangerous situation is likely to develop without intervention. This creates a danger that the operator may pay less attention to the alarms, and alarms may be mistakenly ignored. Therefore, the proposal to reduce alarms to just those that require intervention seems appealing. The three basic approaches to alarm reduction are filtering, conditioning and analysis (Goodstein, 1985). *Filtering* systems use logical rules to reduce active alarms in plant transients, e.g. only display the alarm if the pump has been in operation for 10 seconds or longer. Other filtering techniques may help to prevent the cascade of alarms by using 'intelligent' alarms that summarise the information. *Conditioning* may involve the introduction of a "hysteresis" around the alarm limit. Thus the introduction of a small time lag would prevent an oscillatory becoming an alarm. Mode-based conditioning may only allow an alarm to be shown in certain operational modes, to prevent alarm flooding in certain system states, e.g. start-up or shut-down. Alarm *analysis* may be considered to be comprised of three stages: pre-processing, analysis and display (Herbert, Jervis & Maples, 1978). The pre-processing stage concerns alarm validation, the analysis stage determines prime- causes and last-up alarms in some plant areas and the display stage presents the results of the analysis. Human factors concerns are whether the analysis should be performed by the human or the machine (Meister, 1989). All of the alarm reduction techniques still retain the basic approach of attempting to capture and display "raw" plant information (Goodstein, 1985). That is to say, they follow the philosophy of 'one measurement - one indication'. Alarm analysis does move away from that to some limited extent, but introduces some further uncertainty into the adequacy of the analysis, i.e. what degree of confidence can the operator have in the output?

Most alarm suppression techniques are successful in reducing the 'head count' of alarms. Williams (1985) suggests that combining suppression techniques (i.e. filtering, conditioning, and analysis would probably reduce the number of alarms initiated during plant incidents by at least 50%, but also acknowledges the difficulties of implementing the suppression regime. These problems aside, a study by Sanquist & Fujita (1989) compared alarm suppression, in an advanced system, with a system without suppression. The advanced display coded annunciator alarm information by colour. Red indicated anomalies that required an operator response. Yellow indicated caution information that required operator monitoring. Green indicated normal status information that required no operator action. No coding was used on the conventional display. In addition, alarm reduction accomplished by a logic scheme reduced the number of alarms by 80%. Their data indicated that there was an increase in workload associated with the advanced display. This was demonstrated in terms of more control actions and a longer time required to bring the situation under control. Sanquist and Fujita optimistically propose that this could be due to more effective diagnosis and operational control. It certainly suggests a possible shift in cognitive emphasis (Wickens & Kessel 1981), but shows that alarm reduction does not produce those kinds of effects expected. Baker, Gertman, Hollnagel, Holmstrom, Marshall & Øwre (1985) investigated, amongst other things, a logical alarm reduction system but were also unable to show that this led to better performance. Paradoxically, alarm reduction also reduces the amount of redundant information that is available to the operator, which might, if it were present, be used to enhance performance under certain circumstances. This is because the apparent redundancy of information may hide its usefulness in keeping the operator abreast of the state of the process and developments therein, as well as aiding the diagnosis task. It appears that alarm reduction involves the operator in more monitoring and searching activities, if performance is to be sustained.

Thus, whilst alarm suppression certainly appears to reduce the number of alarms present, this 'head count' is not the only criterion for success. The reduction of alarms is only a success if it leads to enhanced operator performance according to a variety of criteria which could include:

- time to diagnosis
- mental workload
- number of control actions
- success of control actions
- quality of diagnoses and control actions
- 'output' performance
- detection rates

From the studies briefly mentioned above (i.e. Baker et al, 1985; Sanquist & Fujita, 1989), it is suggested that whilst the 'head count' is down, the other criteria are not successful, and in some cases appear blatantly unsuccessful. More recent research by Hogg et al (1995) has shown that alarm lists are not effective in enhancing operators' situational awareness during the initial phases of a disturbance. Therefore, one might argue, is there anything to be gained by reducing the amount of information provided? Development of a logic-based alarm reduction system as described by Cortes (1991) claims possible benefits such as improved productivity, reduced process down time, reduced operator stress and lower control room manning. However, these claims have yet to be validated.

ALARM REDUCTION STUDY

The study considers two factors at issue: the ratio of alarm to non-alarm information and the rate at which information is presented. Often these two factors are intertwined. By reducing the non-alarm information the effect is to reduce simultaneously the rate at which information is presented. For example if 60 alarms are presented in a minute, the rate is one per second. If alarm reduction techniques halve the number of alarms then the rate of presentation will have to become one alarm every two seconds. The experiment conducted attempts to determine which of these two factors makes the difference, the rate of information presentation or the ratio of alarm to non-alarm information. Consideration of the literature led to the expectation that increasing the ratio of alarm to non-alarm information (as could reasonably be expected by introducing alarm reduction techniques) would have no effect on performance, but reducing the rate of presentation would.

METHOD

Participants

Forty five people participated in this study. All were treated according to the British Psychological Societies guidelines governing experimental studies involving human volunteers. The experimental participants were aged between 18 and 45 years.

Design

The participants were randomly assigned to one of nine experimental groups, 5 participants in each group. (see table 1). This was a between-subjects study comprising two factors: temporal rate and ratio of alarms. A completely randomised factorial design was chosen to eliminate the possibility of order and practice effects. The temporal factor comprised three conditions: one alarm per 1 second, one alarm per 4 seconds, and one alarm per 8 seconds. The ratio factor comprised three conditions: 2 percent of target alarms, 6 percent of target alarms, and 10 percent of target alarms. Manipulation of the factors was conducted to see which one was important for alarm detection performance. The main dependent variable was alarm detection accuracy.

TABLE ONE ABOUT HERE

Table 1. Experimental design.

Equipment
The experimental task was written in SuperCard and was run on Macintosh II. participants were required to use the mouse (for the primary task) and two keys marked "S" for same and "D" for different (for the secondary task)

Task
Participants were required to attend to a primary and secondary task. The primary task required them to identify if the message presented in a scrolling text display was one of the target 'alarm' messages, or a non-target message (see figure 1). This was a matching and categorisation task. To the right of the screen a number of alarm messages were presented. To the left of the screen four target buttons and a non-target button were shown. The participants task was to categorise the top, highlighted alarm to the right of the screen as either one of the targets, or as a non-target. This was achieved by moving the cursor by mouse control to the appropriate button and clicking the mouse control. When the primary task allowed, participants were required to make 'same'/'different' judgements about a series of paired figures in different axes of rotation similar to the Shepard & Metzler (1971) task, using two keys to respond. The task was to decide if the figure on the left matched the figure on the right, although it had been rotated. Two keys on the keyboard had been labelled 'same' and different'. After pressing one of these a new rotated figure stimulus was presented.

FIGURE ONE ABOUT HERE

Figure 1. Screen shot of task showing alarm list (right), participant response buttons (top left) and secondary, rotated figures, task (bottom left).

Whilst it is accepted that this task is somewhat artificial, we do not find people performing the task in exactly this form in the 'real world', it can be argued that it contains some necessary elements of detecting and classifying alarm information, albeit in an abstracted form. Laboratory research of this nature enables experimental variables to be presented, controlled and manipulated in a manner that is not possible in more naturalistic environments. This reduces the likelihood of confounding variables affecting the results but can lead to people questioning the validity of the research. There has to be some synergy between laboratory and applied research for progress to be made, each tackling issues developed by one another. Not all research has to, or can, be undertaken in the field with real personnel. In fact this experiment expressly chose to use non-control room personnel on the basis that the task was simple enough to learn within a few minutes and that basic psychological processes, common to all people, were under investigation.

Procedure

On volunteering to participate in the experiment, participants were instructed to read on-screen instructions. These instructions told them about the nature of the two tasks they would be presented with. The alarm detection tasks required the participants to categorise the alarms at the top of the scrolling screen by clicking the mouse button on one of five buttons to indicate if it was a target, or non-target, message. They then had an opportunity to practice this task. Following this the secondary task was explained to them, which again they had a chance to practice. Participants were instructed that the primary task (the alarm detection and categorisation task) should be given priority at all times. When they were sure that they understood the task they were allowed to continue with the main experimental phase, where both tasks were presented simultaneously. On completion of the experimental phase, participants were thanked for their time and involvement in the study.

Measurement

Performance of the participants in both the primary and secondary tasks was measured. Data from the primary task were classified in signal detection terms into: hits, misses, false alarms and correct rejections as shown in table 2.

TABLE TWO ABOUT HERE

Table 2. Four outcomes of signal detection theory

The data were transformed into a index of detectability (Davies & Parasuraman, 1982) : p(A). The following formula shows the transformation explicitly:

$p(A) = ((0.5) + (Y-X) \times (1+Y-X)) / (4xYx(1-X))$

where $Y = H/s$ and $X = F/n$

n = the number of non-target events

s = the number of targets

H = the number of hits

F = the number of false alarms.

The transformation was necessary because different volumes of alarm information were presented in the different conditions as table 3 illustrates.

TABLE THREE ABOUT HERE

Table 3. Number of targets in each condition (in bold text in cells) compared to total stimuli (in plain text in cells)

Therefore p(A) represents an index of detectability (Davies & Parasuraman, 1982) which was used as a measure of the participants' performance in response to the targets embedded within the non target information. Data from the secondary tasks were collected, this included response time and errors.

Analysis

The temporal and ratio data from the primary task were analysed in a two factor ANOVA. The reaction time data from the secondary tasks were also analysed by ANOVA.

RESULTS

The results for the alarm detection task (i.e. the primary task) show that there was not a statistically significant effect for ratio (i.e. 2%, 6%, and 10%) of target to non-target information ($F_{2,36}$= 0.769, p = not significant). This means that reducing the number of non-target alarms, as might occur through alarm reduction strategies, did not improve target alarm detection performance. There was, however, a statistically significant effect for the rate (i.e. 1 second, 4 seconds, and 8 seconds) at which alarms were presented ($F_{2,36}$ = 3.387, $p<0.05$). This means that if people are required to categorise alarms in real-time, detection performance improves as the rate of alarms decreases. The temporal effects were further analysed by Scheffé's F test for post hoc analyses. The results show that participants' target detection performance in the four and eight-second conditions were superior to the one-second condition ($p<0.05$). This is illustrated in figure 2, which shows that p(A) was significantly lower in the one-second condition. There were no statistical differences between the four and eight-second conditions. No interaction between the ratio and temporal factors was found ($F_{2,36}$= 0.016, p = not significant).

FIGURE TWO ABOUT HERE

Figure 2. Histogram of mean sensitivity for each of the temporal conditions.

The results for the secondary task showed that there were no effects of either ratio ($F_{2,36}$= 1.431, p = not significant) or temporal ($F_{2,36}$= 0.369, p = not significant) probability on performance. Nor were there any interaction effects ($F_{2,36}$= 0.36, p = not significant). This means that participants performed the secondary task in a similar manner, despite the differences that the primary task made on them. The results suggest that workload on the secondary task was held at a constant rate and that the variation in performance was observed on the primary task. This could be analogous to a human supervisory control environment where the secondary task takes the role of a general spatial monitoring task and the alarm handling task demands priority, as in the occurrence of a disturbance. This provides some contextual reference for interpretation of the results.

In summary, the results from the experiment reported here indicate that the ratio of alarm to non-alarm information is not necessarily important to detection performance, but the rate at which it is presented, is. It would be difficult to determine an absolute rate of presentation to optimise performance because there are so many influencing variables, such as: type of information presented, context, other demands, knowledge and skill of the human operator, and so on.

DISCUSSION

This study does show that below a certain presentation rate, the quantity of non-alarm information does not impair detection of alarms. This leads to the suggestion that a large reduction in non-alarm information impairs performance (as shown by Sanquist & Fujita, 1989) and a relatively small reduction makes no difference at all. These findings are largely supported by similar studies in the field of vigilance (Mackworth, 1970; Davies & Parasuraman, 1982; Warm, 1984). This leads to the supposition that attention would be better directed to other aspects of alarm system design to improve performance.

The face validity comes from the fact that the main psychological features are abstracted from the real world environment, for example the spatial rotation task was intended to represent a real world background spatial task which had the same demand on the operator under all conditions, such as continuously monitoring a power plant. To some extent this was validated by virtue of the fact that there was no statistical differences between performance in the task under any of the experimental conditions. Spatial tasks have been used in this way in other experimental studies (e.g. Baber, 1991). Modern alarm systems tend to comprise scrolling text displays (Stanton and Baber 1995). The alarm messages used for the alarm handling tasks were taken from a real power plant, e.g ."TURBINE CONTROL OIL PS FAULTY." Alarm information is information which demands some operator intervention. Unfortunately, in most systems this information is embedded amongst other "non alarm" information. This situation is replicated in this experimental task. Hollywell and Marshall (1994) used a much more realistic simulation and got roughly the same findings, but the study was not so well controlled in experimental terms (i.e. the manipulation of ratio and temporal rate). Despite this, they argue that a rate of 15 messages per minute were comfortable for control room operators, provided that they has no other tasks to perform. Temporal presentation rate has become a much bigger problem with increased use of scrolling text displays. Previously a designer of alarm systems would have had to prioritise and pre select which alarms to display on an annunciator board. With information technology the designer has the capacity to provide a far larger quantity of information. Given the technological focus of many alarm designers, there is a tendency for them to present the operator with all the available information, believing more to be inherently better. This confronts the operator with a far higher temporal rate of presentation, which they must sort and prioritise.

There are two conflicting theories of dealing with information overload, perceptual dimming and perceptual narrowing. The perceptual narrowing hypothesis would lead us to believe that people would concentrate on the alarm handling at the expense of the spatial handling task in the conditions of information overload. An alternative theory posits perceptual dimming, this leads to the hypothesis that under conditions of information overload performance more generally would be reduced, but you would not see a focus on one aspect of the task. However the experimental results show that performance of the background task remains unaffected by the rate of delivery of alarm information. This leads us to discount the perceptual narrowing hypothesis and tentatively support the perceptual dimming hypothesis.

It is likely that the absolute number of alarms presented is a side-track of the main issue, which for the purpose of human factors is: can the operator manage the process efficiently and effectively? Therefore, the presence of a large number of alarms may not be a problem if the operator can make sense of them and they do not interfere with the task. The operators themselves may have alarm sifting heuristics, as suggested by Hollywell & Marshall (1994). Such heuristics may include scanning down the list, seeking contextual information and searching for transient-dependent sequences. A substantial influence on the successful management of the incident will be how that information is represented to the operator. Therefore, it has been suggested that a more appropriate solution may involve more effort in: the initial definition of alarms (Usher, 1994), improved methods of presentation (Stanton & Stammers, 1998) and the development of advanced support systems (Williams, 1985). In addressing the question of what to alarm, one should consider to whom the information would be useful. Alarms that are of use to the engineer are not necessarily going to be useful to the operator, and vice versa. Typically, these are mixed within the same system, providing the operator with a lot of irrelevant information that could mask more important alarms. Similarly, defining thresholds to trigger alarms requires careful fine tuning. Unfortunately, plant commissioning can be a hurried process, leaving the operator with many 'false' alarms (Bliss, 1995). Presentation of the information may be largely dictated by technological capability rather than human performance. The introduction of information technology into the control room has not always gone hand-in-hand with improved task performance (Stanton & Baber, 1995).

Research into alarm presentation methods has gone on to demonstrate the relative benefits of different alarm media, such as speech versus text (Stanton & Baber, 1997), comparison of annunciator, mimic and text displays (Stanton & Stammers, 1998), and representative versus abstract auditory displays (Stanton & Edworthy, 1997). One group of researchers demonstrated the effects of combining alarm media (Selcon et al, 1995). All of these studies have been conducted independently of alarm presentation rate. Stanton & Baber (1995) argue that rather than consideration of single messages, it is the alarm gestalt (i.e. the cumulative alarm information) that enables operators to diagnose problems. Provided that the operator has sufficient opportunity to review the alarm information and time to explore alternative courses of action, presentation rate is unlikely to be of much consequence. This only becomes a real issue when the alarm system are poorly conceived, such as scrolling text displays which force operators to try and deal with the information in real-time. Stanton & Baber (1995) argue very strongly against the use of such systems, on the basis of a psychological model of human alarm handling, to propose that parallel, rather than sequential displays, would optimise alarm detection performance. These feelings are echoed by other researchers (e.g. Rauterberg, 1999), who argue that the design and interpretation of the information in context is superior to the deployment of alarm reduction techniques.

CONCLUSIONS

In the design of alarms one should consider who is going to use that information - is it there for the maintenance engineer, the operator, the incident investigator, or the designer of the system. Often different alarms are mixed within the same system, providing the operator with a great deal of irrelevant information that could mask more important alarms. Defining the alarm thresholds is also important. This is borne against the background where in a major incident, e.g Three Mile Island, alarm presentation rates can be somewhere between fifty and three hundred alarms per minute. This experiment clearly shows the limitations of scrolling text displays.

Rather than have things spatially organised, as they were on annunciator panels alarms are presented sequentially, in time. An alarm designer should determine which alarms to include through an understanding of the task that the operator is trying to perform, and the information required to perform that task optimally.

REFERENCES

Andow, P. K. (1983) Improvement of process alarm systems. Eurochem '83.Chemical Engineering Today: The Challenge of Change. (pp 199-205) Oxford: Pergamon Press.

Andow, P. K. & Lees, F. P. (1974) Process plant systems: general considerations . In: Buschmann, C. H. (ed) Loss Prevention and Safety Promotion in the Process Industries. (pp. 299-306) Amsterdam: Elsevier.

Baber, C. (1991) Speech Technology in control Rooms: A Human Factors Perspective. New York: Ellis Horwood.

Baker, S.; Gertman, D.; Hollnagel, E; Holstom, C; Marshall, E & Øwre, F. (1985) An experimental comparison of three computer-based alarm systems: results and conclusions. Norway: Halden Reactor Project (HWR-142).

Bliss. J. P.; Gilson, R. D. & Deaton, J. E. (1995) Human probability matching behaviour in response to alarms of varying reliability. Ergonomics, 38 (11), 2300-2312.

Combs, R. B.; Aghazadeh, F. Ergonomically designed chemical plant control room.. In: Aghazadeh, F. (ed) Trends in Ergonomics/Human Factors V. (pp 357-364) Amsterdam: North-Holland.

Cortes, E. (1991) Logical reduction of alarms (LORA) - an alarm handling software package. IEE Computing & Control Division Colloquium on condition Monitoring for Fault Diagnosis. (IEE Digest No. 1991/156) London: IEE.

Davies, D. R. & Parasuraman, R. (1982) The Psychology of Vigilance. London: Academic Press.

Goodstein, L. P. (1985) Functional Alarming and Information Retrieval. (RisØ-M-2511-18) Denmark: RisØ National Laboratory.

Hale, A. R. & Glendon, A. I. (1987) Individual Behaviour in the Control of Danger. Amsterdam: Elsevier.

Herbert, M. R.; Jervis, M. W. & Maples, G. C. (1978) Past CEGB Experience with Alarm Analysis Systems and Requirements for Future Systems. (RE/LM CONT/063) Letherhead: CEGB.

Hickling, E. (1992) Personal communication.

Hoenig, G.; Umbers, I. G.; Andow, P. K. (1982) Computer based alarm systems. IEE Conference: Trends in On-line Computer Control Systems. London: IEE.

Hogg, D. N.; FollesØ, K.; Strand-Volden, F. & Torralba, B. (1995) Development of a situational awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms. Ergonomics, 38 (11) 2394-2413.

Hollywell, P. D. & Marshall, E. C. (1994) An experiment to support the design of VDU-based alarm lists for power plant operators. In: N. A. Stanton (ed) Human Factors in Alarm Design. (pp 31-44) London: Taylor & Francis.

Kortlandt, D. & Kragt, H. (1980) Process alarm systems as a monitoring tool for the operator. *Proceedings of the 3rd International Symposium on Loss Prevention and Safety Promotion in the Process Industries*. September 15-19, Basle, Switzerland, Vol 1: 10, 804-810.

Kragt, H. & Bonten, J. (1983) Evaluation of a conventional process alarm system in a fertilizer plant. IEEE Transactions on Systems, Man and Cybernetics, 13 (4), 589-600.

Mackworth, J. F. (1970) Vigilance and Attention. Harmondsworth: Penguin.

Miester, D. (1989) Conceptual Aspects of Human Factors. Baltimore: John Hopkins University Press.

Rauterberg, M. (1999) Different effects of auditory feedback in man-machine interfaces. In: N. A. Stanton & J. Edworthy (eds) Human Factors in auditory Warnings. Aldershot: Ashgate.

Sanquist, T. F. & Fujita, Y. (1989) Protocol analysis and action classification in the evaluation of an advanced annunciator system design. Proceedings of the Human Factors Society 33rd Annual Meeting. October 16-20, Denver, Colorado. Vol 2, 1064-1067.

Selcon, S. J.; Taylor, R. M. & McKenna, F. P. (1995) Integrating multiple information sources: using redundancy in the design of warnings. Ergonomics, 38 (11), 2362-2370.

Shepard, R. N. & Metzler, J. (1971) Mental rotation of three-dimensional objects. Science, 171, 701-703.

Singleton, W. T. (1989) The Mind at Work. Cambridge: Cambridge University Press.

Sorkin, R. D. (1989) Why are people turning off our alarms? Human Factors Bulletin, 32, 3-4.

Stanton, N. A. & Baber, C. (1995) Alarm initiated activities: an analysis of alarm handling by operators using text-based alarm systems in supervisory control systems. Ergonomics, 38 (11), 2414-2431.

Stanton, N. A. & Baber, C. (1997) Comparing Verbal Media for Alarm Handling: speech versus textual displays. Ergonomics, 40 (11), 1240-1254.

Stanton, N. A. & Edworthy J. (1998) Auditory affordances in the Intensive Treatment Unit. Applied Ergonomics, 29 (5) 389-394

Stanton, N. A. & Stammers, R. B. (1998) Alarm Initiated Activites: matching formats to tasks. International Journal of Cognitive Ergonomics, 2 (4), 331-348.

Usher, D. (1994) The alarm matrix. In: N. A. Stanton (ed) Human Factors in Alarm Design. (pp 139-145) London: Taylor & Francis.

Warm, J. S. (1984) Sustained Attention and Human Performance. Chichester: Wiley.

Wickens, C. D. & Kessel, C. (1981) Failure detection in dynamic systems. In: J. Rasmussen & W. B. Rouse (eds) Human Detection and Diagnosis of Systems Failures. New York: Plenum Press.

Williams, G. H. (1985) Alarm Loading at Hinkley Point 'B' Power Station - Potential and Suitability of Alarm Suppression. (TPRL/L/2841/N85) Leatherhead: CEGB .

Woods, D. D.; O'Brien, J. F. & Hanes, L. F. Human factors challenges in process control: the case of nuclear power plants. In: G. Salvendy (ed) Handbook of Human Factors. New York: Wiley.

Zwaga, H. J. G. & Veldkamp, M. (1984) Evaluation of integrated control and supervision systems. I.Chem.E.Symposium Series: 90, (pp 133-146). Oxford: Pergamon Press.

Table 1. Experimental design.

Table 2. Four outcomes of signal detection theory

Table 3. Number of targets in each condition (in bold text in cells) compared to total stimuli (in plain text in cells)

Figure 2. Histogram of mean sensitivity for each of the temporal conditions.